Honors Theses at the University of Iowa

Fall 2019

Sensitivity of fine-grained phonetic variation in children who use cochlear implants

Abigail Simon

SENSITIVITY OF FINE-GRAINED PHONETIC VARIATION IN CHILDREN WHO USE COCHLEAR IMPLANTS

by

Abigail Simon

A thesis submitted in partial fulfillment of the requirements
for graduation with Honors in the Speech Pathology and Audiology

McMurray, Bob
Thesis Mentor

Fall 2019

All requirements for graduation with Honors in the
Speech Pathology and Audiology have been completed.

Wu, Yu-Hsiang
Speech Pathology and Audiology Honors Advisor

College of Liberal Arts and Sciences
The University of Iowa
Iowa City, Iowa

SENSITIVITY OF FINE-GRAINED PHONETIC VARIATION IN CHILDREN WHO USE COCHLEAR IMPLANTS

By
Abigail Simon

A thesis submitted in partial fulfillment of the requirements for graduation with Honors in the Department of Communication Sciences and Disorders

Dr. Bob McMurray
Thesis Mentor

Fall 2019

ABSTRACT

Previous studies have found that post-lingually deaf adults with cochlear implants (CIs) generally have good phonemic categorization abilities, but performance can be fragile. Few studies have examined phonemic categorization in pre-lingually deafened children with CIs. This study asks if children who use CIs perceive fine-grained acoustic differences within a category. Next, assuming differences are perceived, we ask if they do so in a manner similar to adults with CIs, and how these patterns change over the course of adolescent development.

This study employed an eye-tracking paradigm to examine perception of voicing and fricative place of articulation in children with CIs. In previous speech categorization experiments, adult CI users typically demonstrate shallower identification functions, which is interpreted as a useful method of adaptation to uncertainty.

Participants ranged in age from 11 to 18 and included 17 CI users and 31 age-matched normal hearing (NH) peers. Children heard a token from either a b/p or s/ʃ continua (eight steps) spanning two words (*bear/pear*, *sip/ship*), and selected the corresponding picture from a screen containing pictures of all four words. Eye movements were monitored while they performed this task to measure how strongly each word was considered over time.

Mouse-click results (phoneme identification) for voicing demonstrated evidence for shallower slopes in CI users for both voicing and fricative continua. With respect to fixations, the CI users showed a less gradient effect of rStep to the competitor. Additionally, the CI users demonstrated minimal activation of competitors. These results suggest that children with CIs may adopt a *wait and see* approach, suppressing competitor activation and waiting to begin lexical access until substantial information has been accumulated.

INTRODUCTION

Accurate speech perception is a complex cognitive task that skilled listeners appear to overcome effortlessly. Speech unfolds rapidly and listeners must simultaneously integrate dozens of brief acoustic cues. While doing this, listeners must account for factors including variability across speakers, co-articulation, speaking rate, and lexical similarity. Within less than half a second per word, skilled listeners will have integrated these cues, accounted for this variability, and mapped the signal onto thousands of potential words in the lexicon. For most listeners with normal hearing, this skill is acquired over the course of development without significant difficulty. However, for listeners with hearing impairments, it can be challenging, especially for children with who are born deaf and use cochlear implants.

Cochlear implants (CI) are an increasingly common remediation approach for individuals with severe-to-profound hearing loss. These devices electrically excite the auditory nerve and allow individuals to hear electrically rather than acoustically. Adults with post-lingual implantation of CIs yield impressive speech perception accuracy. However, this accuracy is fragile and breaks down in noise (Fu et al., 1998).

One of the challenges faced by children with CIs is that they have to acquire the phonemes and lexicon of their native language even as their input is highly degraded. Further, children who use CIs lack any access to auditory information until after activation of their CI, which may create a delayed developmental period. Peripheral processes that allow listeners with normal hearing (NH) to separate target sounds and background noise are impaired due to signal-processing limitations. As a result of these challenges children with CIs likely do not have phonemic categories that are as well formed as peers with NH (Geers & Hayes, 2011). This study focuses on two phonemic distinctions (/b/ vs. /p/ and /ʃ/ vs. /s/) and asks if children with pre-lingual CIs are sensitive to fine-grained phonetic variation within these categories.

The Development of Speech Categorization

One early developmental problem faced by typically developing (TD) children is identifying which sound categories are relevant for their native language. For example, native English-speaking children must learn to distinguish /d/ and /t/, but do not need to learn the difference between a /t/ and dentalized /t̪/; Hindi-speaking infants must learn both. Converging evidence on this issue demonstrates that infants younger than six months can discriminate most speech contrasts across world languages (Werker & Tees, 1984). Then, by approximately 12 months, this discriminatory ability narrows to strictly the native language (Werker & Polka, 1993; Kuhl, Stevens, Deguchi, Kiritani, & Iverson, 2006).

This relatively rapid development of native language discrimination suggests the importance of perceptual learning, based on the acoustic signal (Jusczyk, 1993; Werker & Curtin, 2005), as infants do not have large vocabularies, speech production skills, or other ways to learn these sound categories. One idea is that infants use a form of statistical learning to identify speech contrasts and are sensitive to the frequency distribution of phonetic cues like Voice Onset Time (VOT) and formant frequencies (Maye et al., 2002). Under this model, speech cues form clusters that reflect phonological structure. Statistical learning is used because these young children only have access to acoustic information. They have yet to develop the ability to map onto additional modalities of language like speaking, reading, and writing.

The Development of Speech Perception in Children who use CIs

Young children implanted with CIs may not fit into this traditional statistical learning framework because they do not receive speech input until a later age and may not encode certain acoustic cues with enough fidelity to engage in this form of statistical learning. Most children do not receive cochlear implants until around 12 months of age due to a variety of factors, including reliable diagnosis, anesthetic risk, and challenges in programming (Cosetti & Roland, 2010). Importantly, by 12 months of age, children have gained more advanced cognitive and social skills that are not available to younger infants. These factors (and others) may positively influence the development of speech perception. On the other hand, speech development is likely to be slower as many acoustic cues important for speech perception are more challenging for listeners with CIs to perceive. For example, the /s/ phoneme is important for English grammar but has a high frequency and low intensity and is not always perceived by CI users. Nonetheless, existing studies suggest there are some speech perception abilities available in young children (but not infants).

Peng and colleagues (2019) focused specifically on perceptual discrimination of phonemes for children ages 2 to 3 years old with early cochlear implantation. They used the Reaching for Sound paradigm to collect behavioral responses and played four contrast pairs with the phonemes /p/, /b/, and /k/ and manipulations of place + voicing, place, voicing, and reduced VOT cue. The results demonstrated that young children with bilateral CIs are able to discriminate phonemic-level consonant contrasts above chance level, but more poorly than NH peers (Peng, Hess, Saffran, Edwards, & Litovsky, 2019). Additionally, chronological age was correlated with improved discrimination performance among children with CIs, suggesting gains with development.

Earlier work conducted by Giezen and colleagues (2010) found that 5-6-year-old children with CIs used spectral cues in the /fu/-/su/ contrast less effectively than children with normal hearing. The children with CIs had poorer phoneme endpoint identification and a shallower classification slope. However, these children used similar cue-weighting patterns as their peers with NH (Giezen, Escudero, & Baker, 2010). Despite relatively poor spectral resolution, children with CIs could use spectral cues to aid speech perception, sometimes with as much accuracy as children with NH.

Dunn and colleagues (2014) took a broader look at speech perception (using standardized measures, rather than fine grained measures such as those of Ping and Giezen) longitudinally. They assessed children's performance over time as a function of age of implantation (comparing those under 2 years of age (younger implanted) and those between 2-3.9 years of age (older implanted). Speech perception was assessed with the Consonant-Nucleus-Consonant and Phonetically Balanced-Kindergarten standardized assessments in sound field in a sound-attenuated booth. The younger-implanted group had higher speech perception scores at 5 years of age, compared with older-implanted children. However, this gap narrowed by age 7, and subsequently there was much slower growth for both groups and they appeared to have relatively similar performance trends (Dunn et al., 2014). Significance of implant age group was significant at 8, 9, 10, and 12 years old. These findings suggest that speech perception as a whole may develop very slowly in this population and may not stabilize until early adolescence.

As a whole, this work suggests that speech categorization skills are developing along a fairly later timecourse in CI users (and that they are not as accurate). This mirrors more recent work on typically developing children.

Speech Perception in Older Children

While infants with TD begin to stabilize speech perception abilities by around one year, their abilities continue to develop and are not entirely adult-like (Nittrouer & Studdert-Kennedy, 1987). For example, studies have demonstrated that young children may not have detailed speech category representations and may rely on different acoustic cues for consonant and vowel discrimination than adults (Walley, 2005). These studies typically use a phoneme decision task in which stimuli span continua of two phonemes (e.g. /b/-/p/) and subjects identify each member of the continuum. Subjects in such a paradigm typically show a sharp transition between the two categories. Developmentally, these studies have demonstrated that children do not categorize phonemic contrasts as consistently as adults and display less sharp identification curves which steepen with age (Slawinski & Fitzgerald, 1998; Hazan & Barrett, 2000; Nittrouer & Studdert-Kennedy, 1987).

Many studies clearly demonstrate that speech perception has a developmental trend, so that, for example, twelve-year-olds categorize more similarly to adults than six-year-olds (Hazan & Barrett, 2000). However, twelve-year-olds are still not, on average, categorizing phonemic contrasts as consistently as adults and have less flexibility in perceptual strategies (Hazan & Barrett, 2000). Overall, speech perception is not a skill that is fully developed in infancy. Rather, this skill develops over the course of childhood (McMurray, Danelz, Rigler & Seedorf, 2018).

Categorical vs. Gradient Speech Perception

Most studies using the speech continuum paradigm assume that discrete phonetic categorization is an ideal outcome. This assumption is based on categorical perception, which was a classic finding that suggested some consonants are perceived solely in terms of their discrete label. In the seminal study of Liberman, Harris, Hoffman, and Griffith (1957), adult listeners were presented with a synthetic continuum of consonant-vowel syllables ranging from /ba/ to /da/ to /ga/ by varying formant onset frequency. Using a three-alternative forced-choice labeling task, participants for the most part identified exemplars along this continuum as falling into a single distinct category. In addition, participants were asked to discriminate adjacent steps along the continuum. Critically, they found that discrimination performance along a continuum was largely predicted by category membership. Adults demonstrated discontinuity and heightened sensitivity to formant transition changes at the region of the phonemic boundary. This was taken as evidence that listeners are only sensitive to the phoneme category, not the continuous cues (Liberman, Harris, Hoffman, and Griffith, 1957). These categorical boundaries were assumed to be beneficial for rapidly discarding information and helping listeners quickly decode words. Under this view, when listeners show a more graded slope (e.g., younger listeners), this is assumed to reflect poorer encoding of the speech signal.

In contrast to this idealized notion of categorical perception, recent evidence suggests that individuals are sensitive to within-category differences. This work demonstrates that acoustic information is continuously encoded and independent of phonological information (Toscano, McMurray, Dennhardt, & Luck, 2010). Moreover, continuous encodings are not mapped discretely onto categories. Rather, fine-grained differences are preserved at the level of phoneme categories (Miller, 1997; Andruski, Blumstein, & Burton, 1994; McMurray, Tanenhaus, & Aslin, 2002). As a result of these findings, category boundaries are assumed to be gradient and probabilistic – rather than discrete. Thus, it is possible that shallower slopes (e.g., in CI users or over development) reflect a more gradient – but not less accurate – mode of perception.

In one of the earliest demonstrations of this, McMurray, Tanenhaus, and Aslin (2002) presented adult listeners with a continuum of sounds spanning two phonemes (e.g. /b/ to /p/) and asked them to select a picture of the corresponding word (e.g. *bear* or *pear*). Consistent with categorical perception, listeners labeled the first few steps as /b/ with a sharp transition to /p/.

Critically, this experiment also tracked eye- movements while the referent was selected, making it a visual world paradigm (VWP) task. The logic of eye-tracking was to measure activation of potential referents under consideration, of which the proportion of fixations can reflect the degree of consideration (Tanenhaus, Spivey-Knowlton, Eberhart, & Sedivy, 1995). This study showed that as VOT approaches the category boundary, participants made an increasing proportion of fixations to the lexical competitor that differed in voicing; these gradient effects held across a range of VOTs. Further, listeners demonstrated gradiency even when the authors restricted analysis to only trials that overly identified /b/.

Within Category Sensitivity in Cochlear Implant Users and Children

There has been no investigation of this kind regarding sensitivity to within-category information in children who use CIs. However, previous experiments have built on this approach to understand how and if within-category differences in VOT and additional cues are perceived by adults with CIs.

Preceding work conducted in our lab used eye-tracking to investigate if post-lingually deafened adults with CIs preserve fine-grained differences, or if they maintain a more categorical approach (McMurray, Farris-Trimble, Seedorff, & Rigler, 2015). The authors used the McMurray et al. (2002) paradigm to test CI and NH listeners. Participants heard tokens from several speech continua for voicing and frication (e.g. *bear/pear and ship/sip*) and selected the referent from a screen containing both endpoints and an unrelated minimal pair. Eye movements were monitored to measure word considerations. The results indicated that in both CI and NH listeners, phoneme categorization is not a process of mapping continuous cues to discrete categories. Rather, listeners who use CIs preserve gradiency as equally as listeners with NH as a way to deal with uncertainty. However, they also adapt to their CI by amplifying competitor activation (more than NH listeners).

The above protocol was also used to investigate potential gradient effects of development of speech perception in children with NH. The results of this study indicated ongoing sharpening of speech categories through age 18 (McMurray, Danelz, Rigler & Seedorff, 2018). Children exhibited increasingly gradient within-category sensitivity with development. This suggests that this sensitivity is what enables more discrete categorizations. Further, it implies that speech development is a protracted process in which sensitivity to within-category detail enables increasingly strong phonetic categories (McMurray, Danelz, Rigler & Seedorff, 2018).

Present Study

Our study built upon prior findings that speech perception continues to develop throughout childhood and post-lingually deafened adults with CIs display similar patterns as adults with NH (McMurray, Danelz, Rigler & Seedorff, 2018; McMurray, Farris-Trimble, Seedorff & Rigler, 2016; McMurray, Munson & Tomblin, 2014).

We used the McMurray et al. (2002) paradigm to test children with CIs in a cross-sectional design comparing two age-groups: 11- to 14 and 15- to 18-year olds. Children heard tokens from several speech continua (e.g. *beach* to *peach* and *sack* to *shack*) and selected the referent from the screen. The screen displayed images of the target, competitor, and an unrelated minimal pair (e.g. *bear/pear* and *sip/ship*). Eye movements were monitored throughout the duration of each trial to measure how strongly each word was considered over time. In addition to the voicing VOT, we analyzed fricative place of articulation (e.g. /s/-/ʃ/). As in previous studies, we analyzed eye movements relative to the child’s own boundary and the response on that trial.

We addressed the following questions: (1) Do children with congenital deafness, now using CIs, perceive fine-grained acoustic differences, (2) do children with CIs adapt to fine-grained differences in a manner similar to adults with CIs, and (3) how do speech perception patterns develop over childhood for children with CIs?

METHODS

Participants

CI users were recruited through the Department of Otolaryngology at the University of Iowa Hospitals and Clinics (UIHC). These individuals were enrolled in a longitudinal project conducted by UIHC researchers. NH participants were recruited via mass e-mails to the University of Iowa community and the UIHC newsletter, in accordance with the university human subject protocols. Some of these NH participants were drawn from the same subjects that were reported by McMurray et al. 2018. All subjects were compensated $15/hour. Participants ranged in age from 11-18 years and included 17 CI users and 31 NH controls, for a total of 48. This large age range was necessary to achieve a sufficient sample of CI users. Additionally, this age range is appropriate given work showing that speech perception develops over the course of this time period (McMurray, Danelz, Rigler & Seedorff, 2018).

For analysis we split this into two age ranges: 11-14-year olds (Young; N=23) and 15-18-year olds (Old; N=25). The Young group was distributed between seven CI users and 16 NH controls. The Old group was distributed between 10 CI users and 15 NH controls. The average age for CI participants was 14.7 years (range: 11-18), with an average post-implant age of 11.35 years (range: 1 to 16). All CI users were congenitally deaf. Average age of NH controls was 14.7 (range: 11-18). Exact ages of all participants were rounded to the nearest whole number and this value was used for age-matching between children with CIs and NH peers; age-matching was conducted at a ratio of about 1 CI to 2 NH.

CI users, displayed in Table 1, represented a variety of device configurations. Thirteen were in an electric-only configuration, of whom eight used a unilateral implant and five used bilateral implants. Four subjects were in acoustic + electric configurations: two used a hybrid CI

Table 1

Characteristics of the CI Group

Subj. #	Age (years)	Etiology	Age at onset of deafness	Age at first implantation^	Implant configuration	Implant manufacturer *
11	12	hereditary	0 years	14.4 months	Biltaeral	C
23	12	congenital	11 months	15 months	Unilateral	C
32	13	no cause noted	0 years	11.6 years^	Hybrid	C
37	15	hereditary	3;4 years	4.7 months	Unilateral	C
38	17	hereditary	0 years	21.5 months	Unilateral	C
58	16	congenital	0 years	19.1 months	Bilateral	C
60	12	hereditary	0 years	15.8 months	Unilateral	C
64	18	no cause noted	<2 years	28 months	Unilateral	C
69	15	no cause noted	0 years	3.1 years	Bimodal	C
72	17	meningitis	1;1 years	17.8 months	Unilateral	C
75	13	spinal meningitis	<7 years	6.9 years^	Bilateral	C
76	13	congenital	6 months	12.1 years^	Hybrid	C
77	18	no cause noted	<3 years	31.2 months	Unilateral	C
78	16	no cause noted	0 years	4.2 years	Bilateral	C
80	16	CMV	10 months	27.3 months	Bilateral	C
84	16	ototoxicity	10 months	19.7 months	Unilateral	C
100	11	congenital	0 years	11.2 months	Bilateral	C
173	13	congenital	0 years	4.7 years^	Bimodal	C

^ used HAs before CIs
* Cochlear Americas Corporation

with hearing aids ipsilateral and contralateral to the implant and two with a bimodal configuration of a hearing aid contralateral to the implant. Children with CIs varied in age at onset of deafness from birth to 7 years, with an average of 13 months of age. Age at first implantation ranged from 11.2 months to 6.9 years, with an average of 43.81 months. Three children who were first implanted at older ages had previously worn hearing aids.

NH participants were native monolingual English speakers and were typically developing (TD) with normal or corrected-to-normal vision according to parent report. A hearing screening was conducted on these participants to verify qualification. 29 passed with better than 25 dB hearing at four frequencies in both ears (500, 1000, 2000, 4000 Hz), thereby meeting the ASHA (1990) guidelines. Two participants failed only at the 500 Hz frequency, at 30 dB. As this threshold is below clinical intervention, the children were retained.

Standardized Assessments

We ran a small battery of assessments of language and nonverbal abilities to quantify individual differences. Average scores for each assessment are displayed in Table 2. We assessed receptive vocabulary using the *Peabody Picture Vocabulary Test* (PPVT-IV). The *Block Design* and *Matrix Reasoning* subtests of the *Wechsler Abbreviated Scale of Intelligence* (WASI-II) were administered to assess nonverbal IQ. These three measures were unavailable for one CI participant (Old group). Expressive vocabulary and word retrieval were measured using the *Expressive Vocabulary Test* (EVT-II). This assessment was unavailable for six NH and two CI participants (six from Young group, two Old) due to time constraints.

Each group had average scores within the normal range on assessments, except for the Old CI group for the PPVT. Both CI groups had lower vocabulary scores on the PPVT and EVT than their peers with NH. However, non-verbal skills were normal for participants with CIs and more closely matched with participants with NH. There was not an effect of age in the CI group, but the older children with NH had lower scores than their younger peers (although both were high).

Design

Table 2

Standardized assessment scores

Group	**N**	**PPVT**	**EVT**	**WASI-BD**	**WASI-MR**
Young CI	7	88.2 (19.9)	92.6 (17.2)	50.8 (10.1)	51.3 (13.9)
Old CI	10	83.6 (19.7)	90.2 (11.9)	48.7 (10.2)	53.9 (11.7)
Young NH	16	114.9 (16.1)	114.1 (16.6)	51.7 (11.4)	45.3 (11.8)
Old NH	15	105.8 (12.3)	107.1 (11.9)	54.8 (9.5)	47.2 (9.2)

Standardized assessment scores organized by age group and category with SD in parenthesis. WASI scores are T-scores (mean=50, SD=10) and EVT scores are Z-scores (mean=100, SD=15).

This experiment used six minimal pairs differing in voicing on a /b/ - /p/ continuum (*beach/peach*, *bear/pear*, *bet/pet*, *bin/pin*, *bug/pug* and *bump/pump*) and six differing in fricative place of articulation on a /ʃ/ - /s / continuum *(sack/shack, save/shave, self/shelf, sip/ship, sock/shock).* Each minimal pair consisted of real words easily portrayed in pictures and recognizable for all participants. For each pair, an eight-step continuum was constructed where step one was the most salient /b/ or /ʃ/ sound and step eight was the most distinct /p/ or /s/ sound. On each trial, participants heard a stimulus and selected the referent from a screen containing both endpoints of the target and two from the other class. Thus, a fricative pair served as unrelated referents for a voiced trial and vice versa. The pairing of these images was randomized to avoid emphasizing the relationship between members of a minimal pair. There was a total of two continua types, 6 continua per type, 8 steps, and 6 repetitions of each step. This resulted in a total of 576 trials.

Stimuli

Auditory stimuli were constructed from recordings of a male native English speaker with a standard Midwest dialect. The same stimuli were used by McMurray et al. 2018.
All recordings were made in a sound attenuated booth. Words were recorded using a carrier phrase (“He said __”) and then edited to be isolated.

The 8-step VOT continua were created via progressive cross-splicing. For the voicing stimuli, one sample from each endpoint was selected. Next, a predetermined duration of material was deleted from the onset of the voiced stimuli and replaced with a corresponding segment of the voiceless stimuli. This procedure was done every 8 milliseconds (msec) at the zero-crossing. This method of cross-splicing created an 8-step continuum from 0-56 msec of VOT for each minimal pair.

The fricative continuums were created by first removing the /s/ and /ʃ/ stimuli and equating them in length by removing material at zero-crossings from the middle of the longer of the two stimuli. Long-term average spectra were taken from each sample to determine their spectral averages. Next, we shifted the /s/ and /ʃ/ to have the same spectral mean and averaged the two spectra to create eight steps. After that, these spectra were shifted in frequency space in 8 linear steps. This was then used as a filter for white noise to create the frication noise. Lastly, we computed the average envelope of each phoneme and imposed it on this noise to create the final fricative.

Visual stimuli were constructed using standard lab procedure. A commercial clipart database was searched for several image options of each word and a focus group of lab members determined the best image for each word. Lab members made selections based on minimal visual distractions and continuity across stimuli. Next, each selected image was edited for clarity with

Photoshop. Lastly, all pictures were given final approval by an investigator experienced with the VWP.

Procedure

After giving informed consent, participants were directed to a sound booth and seated in front of a 17” computer monitor, standard keyboard and mouse. Auditory stimuli were presented over loudspeakers placed to the left and right of the monitor. Initial volume of the loudspeakers was 60 dB, but participants were informed that volume could be adjusted during a brief trial period. A padded chin rest was placed in front of the monitor and a research assistant adjusted it to be comfortable for the participant. Next, the research assistant calibrated the eye tracker and provided verbal instructions for the experiment. All participants were informed they could take a break after every 32 trials.

Before beginning the testing phase, participants completed a practice round to better understand the experiment and adjust speaker volume. After this, participants completed a second practice round to gain familiarity with all visual and auditory stimuli. Participants advanced through the visual images paired with the appropriate written word and auditory recording. Before beginning the testing phase, participants had one last opportunity to ask any clarifying questions.

During the testing phase, participants saw four 300-pixel images (50 pixels from each screen corner) and a red dot in the middle of the screen. After 500 msec elapsed, the dot turned blue and the participant clicked on the dot to prompt auditory stimulus. After hearing the word, the participant used the mouse to click on the picture of the word they heard. All participants were informed to take all time necessary and perform the task as naturally as possible. Mouse clicks and eye tracking data were collected during this phase. After completing the task, all participants were debriefed on the experiment and able to ask questions.

Eye movement recording and analysis

Eye movements were recorded using a SR Research Eyelink 1000 desktop-mounted eye-tracker with standard 9-point calibration. A drift correction was done every 32 trials in order to account for any natural shifts and to maintain quality calibration. If the participant failed a drift correct, a research assistant intervened and recalibrated the eye-tracker.

Recorded eye movements were classified as saccades, fixations, and blinks. These events were categorized into "looks," which began at saccade onset and ended at end of fixation (McMurray et al., 2008; McMurray et al., 2002). When converting looks to objects, the regions of interest of the objects were extended by 100 pixels to account for noise in the eye track.

RESULTS

We report two primary analyses: mouse-clicks and eye-movements. The mouse-click data is analogous to traditional phoneme identification measures. Eye-movement data enables us to measure sensitivity of fine-grained distinctions within a category. In both analyses, our goal was to assess the relationship between the phonetic cue (voicing or frication), type of listener (NH or CI) and age (young vs. old).

Mouse-click Analysis

Stop Voicing. Trials in which a participant selected a non-b/p word were eliminated. This eliminated 21 trials from NH listeners and 28 from the CI users, out of a total of 13,535 trials. Figure 1A shows the raw proportion of /p/ responses as a function of VOT step (1-8) and

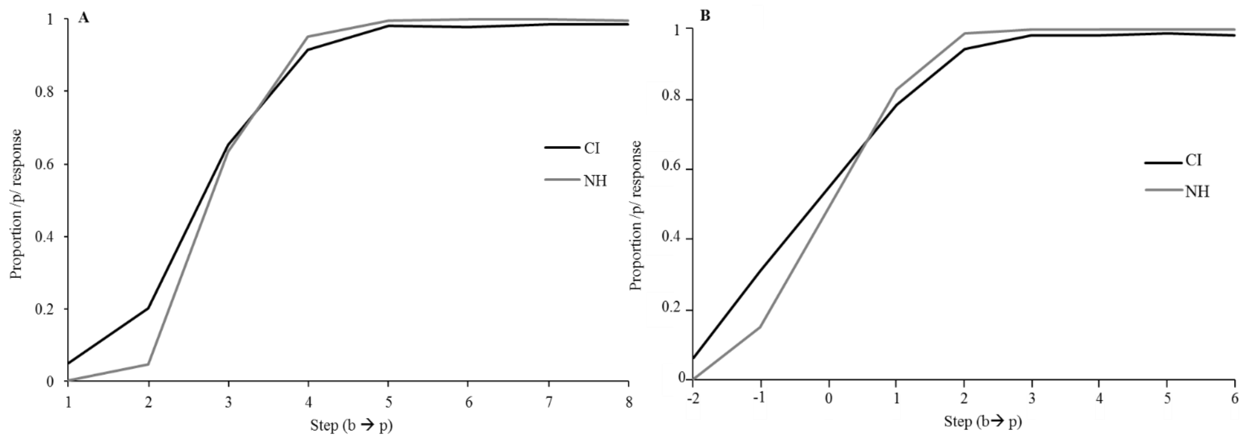

Fig 1. Mouse click (identification) of the voicing continuum. A, as a function of absolute voicing step. B, as a function of step computed relative to each subject x continuum category boundary (rStep).

listening type (CI or NH). Overall differences appear relatively small; the most apparent difference was demonstrated between the slope of NH listeners and CI users, where the CI users showed a shallower slope.

However, this visual difference could be due to an averaging artifact; CI users could demonstrate equally sharp categorization but have more variation in the location of their boundaries across individuals. This would then appear as a shallower slope when averaged across listeners. As in prior studies (McMurray, Danelz, Rigler & Seedorff, 2018; McMurray, Farris-Trimble, Seedorff & Rigler, 2016), to account for this possibility, we computed each participant's own boundary and recomputed the continuum step relative to that to compute relative step (rStep). In this scale, each participants own boundary is at an rStep of 0, making -1 one step toward the voiced end of the continuum and +1 one step toward the voiceless end.

To calculate rStep, we fit a four-parameter logistic function to each subject's data and used the crossover as the subjects boundary (McMurray, 2019). Figure 1B displays the same data, but as a function of rStep (to account for boundary variability). This data demonstrates a clear difference in slope between NH listeners and CI users.

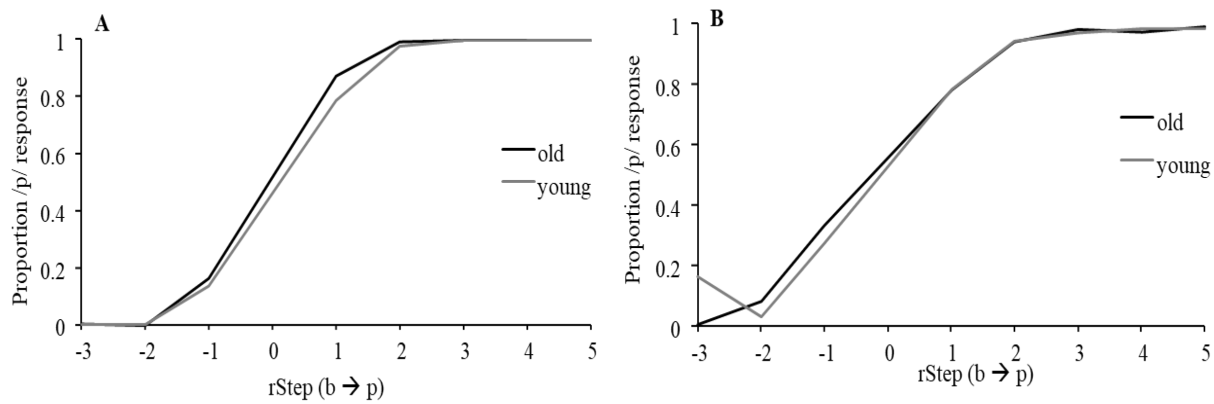

Fig 2. Mouse click (identification) of the voicing continuum as a function of rStep separated by age group. A) for NH listeners. B) for CI users.

Figure 2 displays the proportion of responses as a function of rStep but broken down by age. Differences were most visually apparent between young and old CI users (Figure 2B). The only exception was that the young CI users never fully categorized the /b/ phoneme at (the most extreme rSteps), while both young and old groups fully categorized the /p/. The NH listeners seemingly demonstrated no differences between age groups and fully categorized both ends of the continuum.

For statistical analyses, we examined the parameters of the logistic function used to estimate the boundaries. This function has four parameters. First, the slope describes the rate of change in proportion to /p/ responses as step number changed. Second, the crossover is the boundary. This was used to compute rStep but not otherwise analyzed. Third, and fourth, this function estimates the maximum and minimum asymptotes. These values were combined into a single measure, amplitude, of their difference. Data were analyzed with a between-subjects ANOVA which assessed two factors: hearing configuration (CI or NH) and age group (young or old).

We ran two separate ANOVAs with slope and amplitude as the dependent variables, respectively. Our first analysis examined slope. We found a significant main effect of hearing configuration on slope (F (1,40) =27.035, *p*<.001). This was because CI users had a shallower slope than NH listeners. There was not a significant main effect of age (F (1,40) = .578, *p*=.452). Additionally, there was not a significant interaction between hearing configuration and age group (F (1,40) = 1.389, *p*=.245).

Regarding amplitude, we found a significant main effect of hearing configuration (F (1,40) = 9.404, *p*=.004). The children with CIs have a smaller amplitude because they don't fully categorize this voicing. There was no significant effect of age group and no significant interaction between configuration and age group (F (1,40) = .308, *p*=.582; F (1,40) = .552, *p*=.462).

Overall, the findings for slope and amplitude both denote significant differences due to hearing configuration and no significant effects of age or an interaction between hearing configuration and age.

Fricative Place of Articulation. Trials in which a participant selected a non-s/ʃ word were eliminated. This eliminated 22 trials from NH listeners and 46 from the CI users, out of a total of 13,546 trials. Again, in Figure 3A, we first display the proportion of /s/ responses as a function of frication step and listening type. Overall differences appeared to be relatively small, with the most noticeable difference being overall slope for type of listener. Additionally, the CI users appear to not fully categorize the /s/ phoneme. To account for a potential averaging artifact, we computed the rStep, which is displayed in Figure 3B. Replotting this data demonstrates again a clear difference in slope.

We further analyzed within-group differences between young and old listeners, as shown in Figure 4. Figure 4A displays the identification curves for each age-group in children with NH and Figure 4B displays the same for CI users. NH children display nearly identical patterns of

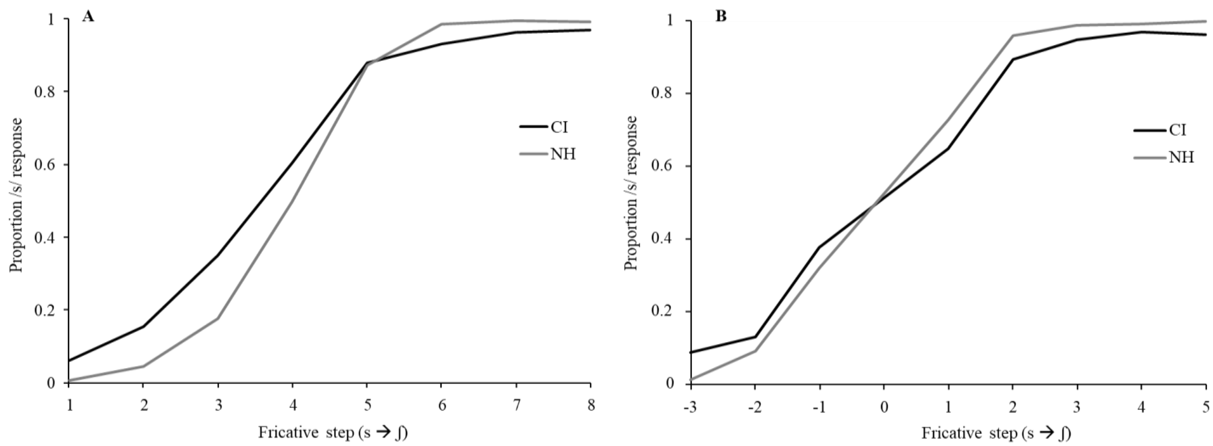

Fig 3. Mouse click (identification) of the voicing continuum. A, as a function of absolute voicing step. B, as a function of step computed relative to each subject x continuum category boundary (rStep).

categorization regardless of age. However, older CI users appear to have a slightly steeper curve when compared to younger CI users.

These data were analyzed with an ANOVA with two factors: hearing configuration (CI or NH) and age group (young or old). The first ANOVA examined slope and did not find a significant effect of configuration (F (1,43) =1.95, *p*=.17). We also did not find a significant effect of age group (F (1,43) =.323, *p*=.572). Lastly, we did not find a significant interaction between age group and configuration (F (1,43) = .111, *p*=.74).

Regarding amplitude, we found a significant main effect of hearing configuration (F (1,43) = .12.84 *p*=.001). We did not find significant effects of age group or of the interaction between age group and configuration (F (1,43) = .072, *p*=.789; F (1,43) = .032, *p*=.86). Overall, the findings for slope demonstrate no significant interactions. For amplitude, hearing configuration is the only statistically significant factor.

In summary, identification results for the voicing continua suggest that CI users have significantly shallower slopes and amplitudes to their identification functions. The children with CIs were fairly good at detecting voicing but were still impaired. Age of the participant,

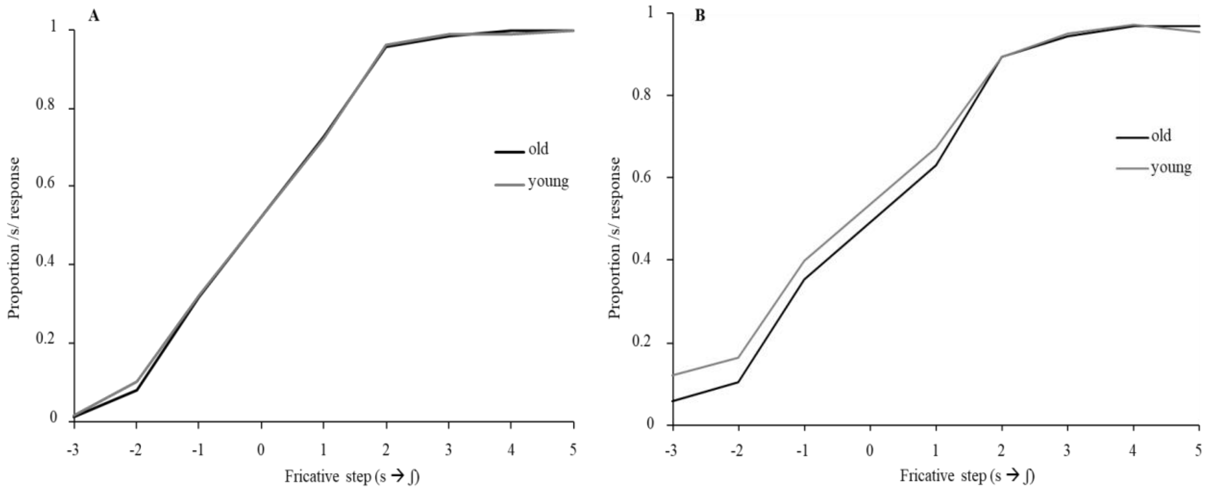

Fig 4. Mouse click (identification) of the voicing continuum separated by age group. A, as a function of rStep for NH listeners. B, as a function of rStep for CI users.

regardless of their hearing configuration, did not significantly influence categorization ability. Results from the fricative continua suggest that CI users have a more significant difference between their absolute categorization points, compared to children with NH. Unlike the voicing continua, CI groups did not demonstrate significantly shallower slopes than their peers with NH.

Eye-movement Analysis

Analysis of eye-movements was accomplished by computing the proportion of looks to each object that the participant made at each competitor at each 4-ms time slice. Figure 5 shows the representative time course of fixations for several conditions at step 1 of the b/p and step 1 of the s/ʃ continuums (unambiguous /b/ or /ʃ/). Figure 5A shows the NH listeners after hearing a clear /b/ token with a VOT of 0 msec. By about 500 msec, these listeners are already fixating on the target (/b/) more than the competitor (/p/) or unrelated (s/ʃ) items. There are few looks to the competitor and unrelated items as the VOT was unambiguous. Figure 5B shows that with the same stimuli, CI users have a slower onset of fixations and maintain these fixations for longer. The fricatives (5C and 5D) show heightened looks to the competitor, but overall similar group differences. The NH listeners demonstrate an earlier peak activation to the competitor and unrelated and fixate to the target by about 750 msec. The CI users maintain active competition until about 1,000 msec.

Our analysis of the eye-movements focused on the degree to which CI and NH listeners are sensitive to differences in VOT or frication within a category. For a set of trials unambiguously heard as a /b/, we sought to determine to what degree listeners are still sensitive to changes in VOT. This allows us to examine sensitivity the residual sensitivity to fine-grained changes. Our primary eye-movement analysis concerned how looks to the competitor (e.g., *beach* when the stimulus was a VOT consistent with *peach)* are modulated by within-category stimuli changes. As in Figure 2, we recoded continuum step in terms of distance from the boundary (rStep) to eliminate variability between participants. We also eliminated all trials in which the participant selected a competing response. For example, if the rVOT was -1 (a /b/ one

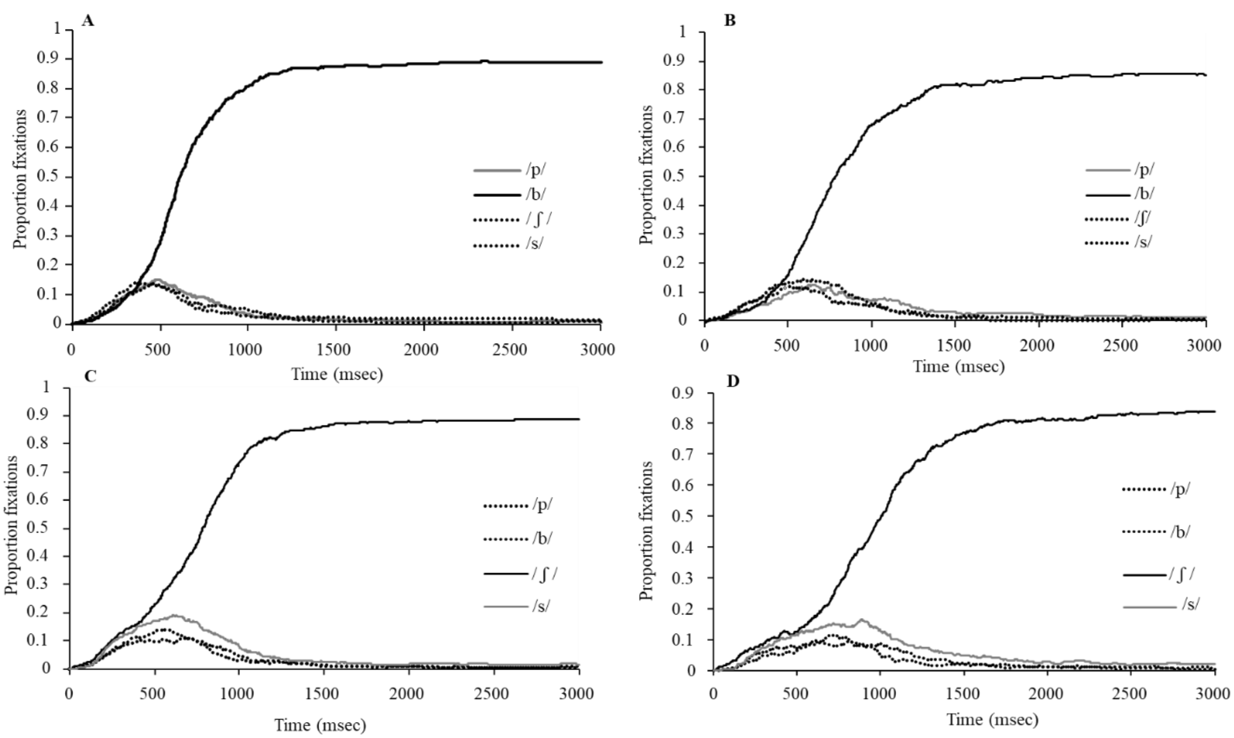

Fig 5. Time course of fixating the target, competitor, and unrelated objects. A, for NH listeners after hearing step 1 of /b/. B, for CI users after hearing step 1 of /b/. C, for NH listeners after hearing step 1 of /s/. D, for CI listeners after hearing step 1 of /s/.

step away from that participant's boundary), we eliminated any trial in which /p/ was the response.

Figure 6 shows representative data from the /b/ side of the stop voicing continua by looks to /b/ when participants heard /p/. It displays fixations to the competitor objects over time as a function of distance from the boundary (rStep). The Young NH listeners (Figure 6A), showed a gradient pattern with more fixations to the competitor for rSteps nearing the boundary (closer to 0). The Old NH listeners (Figure 6B) showed a similar gradient pattern, but steeper. CI listeners (Young in 6C and Old in 6D) appear to have no meaningful gradient looks when younger but develop to have a gradiency effect when older. Additionally, older CI users have longer competitor fixations compared to the older NH listeners.

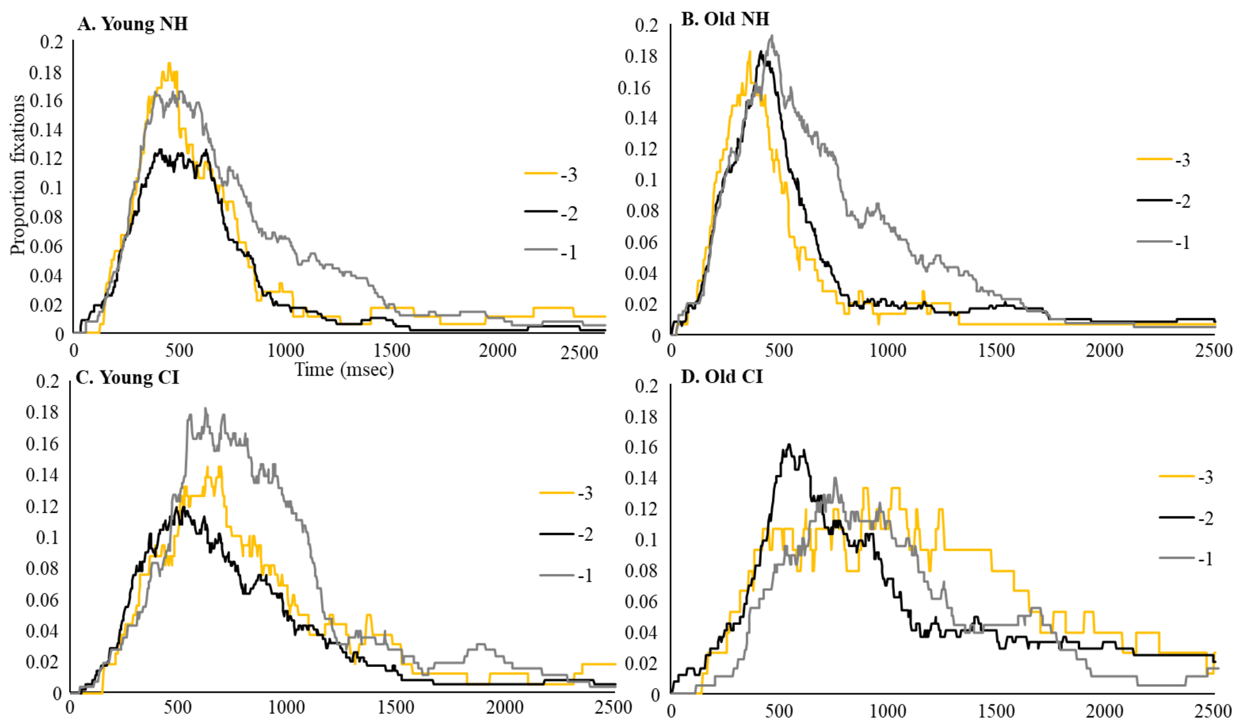

Fig 6. Looks to the competitor as a function of time and relative step for the /b/ side of the voicing continua. For example, -1 is one step away from the boundary, -2 is two steps away from the boundary, so forth.

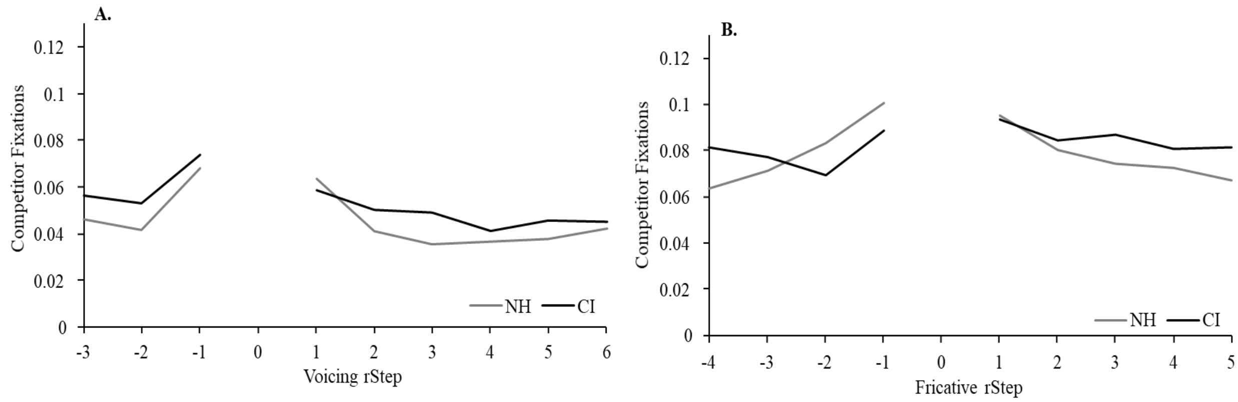

Fig 7. Looks to the competitor (area under curve) as function of rStep and listener group for (A) voicing continua and (B) fricative continua.

To identify group differences, we averaged across time but within group and rStep to compute the number of fixations to the competitors under the curves; the results of this process are shown in Figure 7. Consistent with prior work, subjects showed a gradient effect of distance from the boundary; as the speech sound approached the boundary, listeners tended to look more at the competitor object (McMurray, 2015).

Stop Voicing. To analyze looks to the target we conducted a repeated measures ANOVA with rStep as a within-subjects factor and hearing group, and age group as between-subjects factors. We first examined looks to /p/ when /b/ was the target (rStep <0), averaged between 300 and 2300 msec. We found a main effect of rStep (F (1,41) = 17.491, *p*=<.001). This was because there were more competitor looks as rStep approached the boundary. There was also no main effect of hearing group or age group (F (1,41) = 1.128, *p*=.294; F (1,41) < .0001, *p*=.980). Lastly, there were no significant interactions of rStep with type of listener or with age group (F (1,41) = 1.511) *p*=.226; F (1,41) = 1.575, *p*=.217).

Next, we examined looks to /b/ when /p/ was the target (rStep >0). Again, we found a significant main effect of rStep (F (4,38) = 4.913, *p*<.001). This is due to more looks to the competitor as rStep approached the boundary. As before, there was no main effect of type of hearing group or age group (F (1,41) = .001, *p*=.41; F (1,41) = .003, *p*=.221). There were also no significant interactions of rStep with hearing group or with age group (F (4,38) = 2.349, *p*=.072; F (4,38) = .898, *p*=.88).

***Fricative Place of Articulation*.** We conducted a repeated measures ANOVA with rStep as the within-subjects factor and hearing group and age group as between-subjects factors for the s/ʃ continua. The dependent variable was average looks to the competitor between 300 and 2300 msec.

We first examined looks to /s/ when the target was /ʃ/ (rStep>0). We found a main effect of rStep (F (1,41) = 845.34, *p*<.001). This is because listeners increase looks to the competitor as rStep approached the boundary. We did not find a main effect of type of hearing group (F (1,41) = 4.453E-5, *p*=.904). However, we did find a main effect of age group (F (1,41) = 2.339, *p*<.001). The interaction of rStep with hearing group and age group was not significant (F (1,41) = .001, *p*=.525).

Next, we examined looks to /ʃ/ when /s/ was the target (rStep<0). We found a main effect of rStep (F (4,38) =3.525, *p*=.015). This is because listeners increase looks to the competitor as rStep approached the boundary. We also found a significant interaction of rStep and hearing group (F (4,28) = 3.754, *p*=.011). This is because children with CIs did not increase their looks to the target, even as rStep approached the boundary, while NH children did (see Figure 8B). There was no main effect of type of hearing group or age group (F (1,41) = .427, *p*=.517; F (1,41) < .0001, *p*=.99).

These analyses showed a less gradient effect of within-category differences on competitor activation for children with CIs. For the voicing continuum, there was a main effect of rStep for looks to both competitors. However, age group and listening configuration did not significantly interact with the rStep. For the fricative continuum, there was again a main effect of rStep for looks to both competitors. The main effect of rStep for both continua indicates gradiency of responses. For competitor looks to /s/, there was a main effect of the age of group of listeners. This is important because it indicates that children in the young- and old-groups had different competitor fixations. For competitor looks to /ʃ/, there was an interaction between rStep and type of listeners. This is important because it means that children with CIs did not increase looks to /ʃ/, even as rStep approached the boundary. For the fricative continua at steps less than 0, children with CIs demonstrate less gradiency because they do not increase eye movements as rStep approaches the boundary. In contrast, their peers with NH do.

GENERAL DISCUSSION

The results of this investigation are revealing about speech perception for children with CIs. With respect to the identification measures, we found evidence for shallower slopes in CI users for the voicing continuum and poorer asymptotic performance for the fricatives. This is interesting because despite the significant differences, the children with CIs still performed well with accuracy above 90% on each continuum. With respect to eye fixations, the CI users show a less gradient effect of rStep to the competitor. Additionally, the CI users have a different pattern of activation of competitors (for example, looks to /ʃ/ when target is /s/).

These results demonstrate that children with CIs do not have gradiency for the /s/ side of the fricative continua. This may also be true for the other continua and sides (and is suggested by

Figure 7). However, the current conclusion is limited by the power of the statistical analysis employed. We plan to conduct a stronger analysis using a mixed effects framework and believe that this will reveal significant interactions for rStep and hearing configuration for the /ʃ/-side of the fricative continuum and the voicing continuum. Thus, the following discussion serves to discuss preliminary results and is based on the notion that results will generalize as expected.

Audiological Factors. While we can draw conclusions for CI users more generally, we cannot postulate on differences as a function of CI configuration. The CI users encompassed a variety of configurations: unilateral, bilateral, hybrid, and bimodal. We do not have a large enough sample size to examine differences in categorization dependent on type of listening configuration. However, this is consistent with previous work regarding the CI population and therefore is not a barrier to the results presented. A major challenge in research regarding CIs is variability among participants, even among those with similar demographic and audiologic profiles (Peterson, Pisoni, & Miyamoto, 2010). Additionally, three of the CI users had access to some auditory acoustic information with hearing aids before implantation. However, this is not a likely a major contributor to the results here because the CI users in this study display large differences from the patterns observed in post-lingually deafened adults with CIs (McMurray et al., 2015). Whereas adults increase looks to the competitor compared to adults with NH, we found that children do not have this gradient response.

Limitations. One limitation in our current analysis is the time course used for eye movements of the fricative continuum. We analyzed the range of 300 to 2300 msec, but similar studies conducted previously have used the range of 600 to 2600 msec. This later timecourse has been selected based on work suggesting that lexical access does not begin until the offset of fricatives (McMurray et al., 2016). In our next analysis, we plan to use this later timecourse. A second limitation in our current analysis is that an ANOVA cannot use the entire stimulus range. This will also be corrected in our next analysis.

Summary. This study demonstrated clear patterns of categorization for pre-lingually deafened children with CIs. For the /s/ side of the fricative continua, children with CIs do not demonstrate gradiency and do not have heightened competition near the boundary. This finding is significant because children with CIs display a pattern much different than adults with post-lingual CIs and more similar to same-age peers with NH.

Adult CI users, like adults with NH, show a gradient effect of rStep on fixations to the competitor for voicing and fricatives (McMurray et al., 2015). These adults preserve fine-grained differences and activate categories in a more graded manner. However, they also show heightened competitor fixations overall – something that was not observed here. Thus, adults with CIs and NH use gradiency by maintaining competing alternatives because it is a useful way to deal with uncertainty.

Children with NH slowly sharpen phoneme categorization over the course of development and become increasingly sensitive to fine-grained detail, but do not have this ability at a younger age (McMurray et al., 2018). They appear to slowly develop gradient, within-category activation of lexical competitors over the course of adolescence. Children with CIs do not demonstrate much gradiency. This brings into question whether it is possible that CI users will develop this fine-grained categorization ability, just at a much later age.

As a result of this study, one method used could be the wait-and-see approach, wherein speech information accumulates and listeners wait to initiate lexical access until sufficient information is available to identify the target word (McMurray, Farris-Trimble & Rigler, 2017). Similar to the prelingually deaf-children of McMurray et al. (2017) the children with CIs in our study demonstrate a wait-and-see approach. As demonstrated by time course of fixations (Figure 5), the CI users have delayed activation of the target and increased activation of competitors. Our finding builds on previous work and demonstrates that earlier-implanted children display the same wait-and-see approach as children implanted at later ages. It is not clear whether this wait-and-see strategy for children with CIs is useful or hinders their ability for lexical access. One possibility is that this strategy is a natural, maladaptive consequence of forcing the lexical system to deal with degraded input. Another possibility is that this strategy is adopted as a kind of coping mechanism for degraded input, wherein attempting to process more immediately would decrease accuracy even lower. These results bring in to question which method of learning children with CIs are employing to deal with degraded input and process speech.

Finally, we might ask about how children gained these phoneme categories to begin with. As discussed in the introduction, statistical learning seems insufficient because it does not account for childhood and adolescent development. It is also likely that children with CIs cannot map onto statistical learning like children with NH. These children lack early auditory access and do not encode certain acoustic cues with enough fidelity.

One possibility is similar to the lexical restructuring hypothesis (Metsala and Walley, 1998) and is that children with pre-lingual CIs learn phonological categories via the auditory contrast of minimal pairs. Under this model, children are exposed to spoken minimal pairs and forced to learn the meaningful differentiations, for example between “beach” and “peach”. A second possibility is that reading instruction, particularly during the early developmental period, fosters phonological differentiation abilities. In reading and spelling curricula, children are taught phonological tasks like rhymes and are taught letter/sound mappings. This may force children to learn phonological categories.

Conclusions. The development of fine-grained speech perception has been widely studied for children with NH, but not for children with CIs. This study demonstrates that children with CIs do not have fine-grained phonetic discrimination. Therefore, these children perceive speech in a manner quite different to adults with post-lingual CIs. This pattern of speech perception may change over the course of development, as significantly demonstrated on the /ʃ/-side of the fricative continuum. Future analyses of these findings will seek to validate that children with CIs do not have gradiency and seek to answer if this changes over the course of development. This study serves to show that more research needs to examine the perception and development of fine-grained information in speech for children with CIs.

ACKNOWLEDGEMENTS

The author thanks Camille Dunn for assistance with CI patient recruitment, and Claire Goodwin, Kristin Rooff, Francis Smith, and Keith Apfelbaum for ongoing help throughout the project. The author also thanks Bob McMurray for incredible mentorship and advice.

This work was supported by NIH Grant # P50DC000242 awarded to Bruce Gantz and Bob McMurray. Additional funding was provided by the Iowa Center for Research by Undergraduates.

REFERENCES

Andruski, J. E., Blumstein, S. E., & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition, 52(3),* 163-187.

Aslin, R.N., Pisoni, D.B., Hennessy, B.L., & Perey, A.J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effect of early experience. *Child Development, 52*, 1135–1145.

Carney, A.E., Widin, G.P., & Viemeister, N.F. (1977). Non-categorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America*, *62,* 961–970.

Clayards, M., Tanenhaus, M.K., Aslin, R.N., et al. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition, 108,* 804-809.

Cooper, R. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology, 6,* 84–107.

Cosetti, M., & Roland, J. (2010). Cochlear Implantation in the Very Young Child: Issues Unique to the Under-1 Population. *Trends in Amplification, 14*(1), 46-57.

Desai, S., Stickney, G., Zeng, F. G. (2008). Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. *Journal of the Acoustical Society of America, 123,* 428-440.

Dunn, C. C., Walker, E. A., Oleson, J. B., Kenworthy, M. I., Van Voorst, T. J., Tomblin, J., . . . Gantz, B. (2014). Longitudinal Speech Perception and Language Performance in Pediatric Cochlear Implant Users: The Effect of Age at Implantation. *Ear and Hearing, 35*(2), 148-160.

Eimas, P.D., Siqueland E.R., Jusczyk P.W., & Vigorito J. (1971). Speech perception in infants. *Science, 209,* 1140–1141.

Farris-Trimble, A., McMurray, B., Cigrand, N., & Tomblin, J. B. (2014). The process of spoken word recognition in the face of signal degradation. *Journal of Experimental Psychology: Human Perception and Performance, 40(1),* 308.

Friesen, L.M., Shannon, R.V., Baskent, D., et al. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *Journal of the Acoustical Society of America*, *110,* 1150-1163.

Fryauf-Bertschy, H., & Tyler, R.S. (1997). Cochlear implants use by prelingually deafened children: The influences of age at implant and length of devise use. *Journal of Speech, Language & Hearing Research, 40*(1), 183.

Fu, Q.J., Shannon, R.V., Wang, X. (1998). Effects of noise and spectral resolution on vowel and constant recognition: Acoustic and electric hearing. *Journal of the Acoustical Society of America, 104,* 3586-3596.

Giezen, Marcel R., Escudero, Paola, & Baker, Anne. (2010). Use of acoustic cues by children with cochlear implants. (Report). *Journal of Speech, Language, and Hearing Research, 53*(6), 1440-1457.

Hamzavi, J., Baumgartner, W.D., Pok, S.M., et al. (2003). Variables affecting speech perception in postlingually deaf adults following cochlear implantation. *Acta Otolaryngol, 123,* 493-498.

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics, 28(4),* 377-396.

Kishon-Rabin, L., Gehtler, I., Taitelbaum, R., Kronenberg, J., Muchnik, C., & Hildesheimer, M. (2002). Development of Speech Perception and Production in Children with Cochlear Implants. *Annals of Otology, Rhinology & Laryngology, 111*, 85-90.

Lane, H., Denny, M., Guenther, F.H., et al. (2007). On the structure of phoneme categories in listeners with cochlear implants. *Journal of Speech, Language, and Hearing Research, 50,* 2-14.

Liberman, A.M., Harris, K.S., Hoffman, H.S., & Griffith, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54,* 358-368.

Massaro, D. W., & Cohen, M. M. (1983). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 9(5),* 753.

Metsala, J.L. & Walley, A.C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J.L. Metsala & L.C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Mahwah, New Jersey: Erlbaum.

McMurray, B. (2019, November 13). Nonlinear Curvefitting for Psycholinguistic (and other) Data. Retrieved from osf.io/4atgv

McMurray, B., & Aslin, R.N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition, 95,* B15-B26.

McMurray, B., Aslin, R. N., Tanenhaus, M. K., Spivey, M. J., & Subik, D. (2008). Gradient sensitivity to within-category variation in words and syllables. *Journal of Experimental Psychology: Human Perception and Performance, 34(6),* 1609.

McMurray, B., Farris-Trimble, A., Seedorff, M., & Rigler, H. (2015). The effect of residual acoustic hearing and adaptation to uncertainty in Cochlear Implant users. *Ear and Hearing*.

McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive psychology, 60(1),* 1-39.

McMurray, B., Tanenhaus, M.K., Aslin, R.N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition, 86,* B33-B42.

Moberly, A.C., Lowenstein, J.H., Tarr, E., et al. (2014). Do adults with cochlear implants rely on different acoustic cues for phoneme perception than adults with normal hearing? *Journal of Speech, Language, and Hearing Research, 57,* 566-582.

Mullennix, J., & Pisoni, D.B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception & Psychophysics, 47,* 379-390.

Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech, Language, and Hearing Research, 30(3),* 319-329.

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R. S. (1989). The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech, Language, and Hearing Research, 32(1),* 120-132.

Peng, Z. E., Hess, C. R., Saffran, J. Y., Edwards, J., & Litovsky, R. (2019). Assessing Fine Grained Speech Discrimination in Young Children with Bilateral Cochlear Implants. *Otology & Neurotology, 40*(3), E191-E197.

Peterson, N.R., Pisoni, D.B., & Miyamoto, R.T. (2010). Cochlear implants and spoken language processing abilities: review and assessment of the literature. *Restorative neurology and neuroscience, 28 2*, 237-50.

Pisoni, D.B., & Lazarus. J.H. (1974) Categorical and noncategorical modes of speech perception along the voicing continuum. *Journal of the Acoustical Society of America, 55,* 328–333.

Schouten, B., Gerrits, E., & van Hessen, A. (2003). The end of categorical perception as we know it. *Speech Communication, 41(1),* 71-80.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268(5217),* 1632-1634.

Torpy, Janet M., & Summerfield, A. Quentin. (2002). Use of vocalic information in the identification of /s/ and /[integral]/ by children with cochlear implants. (Ear & Hearing). *JAMA, The Journal of the American Medical Association, 287*(20), 2632.

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous perception and graded categorization electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychological Science*, *21(10)*, 1532-1540.

Walley, A.C. (2005). Speech perception in childhood. In *Handbook of Speech Perception,* Edited by: Pisoni, D.B. and Remez, R.E. 449–468. Oxford: Blackwell.

Werker, J. F., & Polka, L. (1993). Developmental changes in speech perception: New challenges and new directions. *Journal of Phonetics, 21,* 83-101.

Werker, J.F., & Tees, R.C., (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7,* 49-63.